DOCUMENT RESUME

ED 298 443 CS 009 288

AUTHOR Harlin, Rebecca; Lipa, Sally
TITLE Literacy Assessment: Does Task Format Make a Difference?
PUB DATE 25 Feb 88
NOTE 23p.; Paper presented at the Annual Meeting of the Eastern Educational Research Association (Miami, FL, February 24-27, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Grade 1; *High Risk Students; Literacy; Primary Education; *Reading Achievement; *Test Reliability; *Test Validity
IDENTIFIERS Emergent Literacy; *Language Delayed; Print Awareness

ABSTRACT

A study examined the effectiveness of both the informal and standardized readiness measures in predicting the literacy development of both normal first grade subjects and high-risk, language-delayed primary children. Subjects, 60 first grade and 27 language-delayed children were given three informal literacy measures, the "Writing Vocabulary Test," the "Concepts About Print Test (CAP)," and the "Sentence Dictation Test," as well as a standardized readiness battery, the "Metropolitan Readiness Test (MRT)," and the "Peabody Picture Vocabulary Test (PPVT)," a language awareness measure. Data were analyzed using Pearson-product-moment correlations, linear, and multiple regression analyses. Results showed that the informal literacy measures were the best predictors of both the reading achievement of the first graders on the "Stanford Achievement Tests," and the high-risk, language-delayed subjects' scores on Clay's "Letter Identification Test." (Two tables of data are included and 20 references are attached.) (RAE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED298443

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Rebecca Harlin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Literacy Assessment: Does Task Format Make A Difference?

Rebecca Harlin

Sally Lipa

State University of New York at Geneseo

Paper presented at the annual meeting of the

Eastern Educational Research Association, Miami, Florida, February 25, 1988

Running Head: LITERACY ASSESSMENT

CS 009288

Literacy Assessment: Does Task Format Make a Difference?

Abstract

A sample of 87 first grade and language-delayed children was given three informal literacy measures, the Writing Vocabulary Test, the Concepts About Print Test (CAP), and the Sentence Dictation Test, as well as a standardized readiness battery, the Metropolitan Readiness Test (MRT), and the Peabody Picture Vocabulary Test (PPVT) a language awareness measure. The effectiveness of both the informal and standardized measures in predicting the reading achievement of the first grade subjects and the high-risk, language-delayed children was examined. When the data were analyzed using Pearson-product-moment correlations, linear, and multiple regression analyses, the informal literacy measures were found to be the best predictors of both the reading achievement of the first graders on the Stanford Achievement Tests, and the high-risk, language-delayed subjects' scores on Clay's Letter Identification Test. Neither the MRT and the PPVT was found to be as effective nor sensitive as the informal literacy tasks. Relationships between subjects' performance on the print awareness measure, the CAP, and the writing assessments, the Writing Vocabulary Test and the Sentence Dictation Test, differed for both groups of subjects. Implications for assessment and instruction of normal and high-risk students are discussed.

Literacy Assessment: Does Task Format Make A Difference?

Introduction

Young children enter school with individual and divergent literacy experiences. Considerable disagreement exists concerning how to best assess children's competencies and utilize the results of instruments designed to tap this knowledge.

Since large groups of children need to be screened prior to formal reading instruction, group standardized tests are presently utilized to differentiate those children in need of preventative strategies from those in need of more formal reading programs. As the age level for school entry becomes younger there is a strong tendency to use the same measures for assessing four year olds as five and six year olds; and the same measures for an identified language delayed population as for a normal population. This policy ignores clear differences in literacy development and the requisite program opportunities that may be essential for younger and high-risk children.

Standardized reading readiness tests, used in a diagnostic manner are known to "drive the curriculum." These tests which assess skills such as auditory discrimination, letter identification, letter-sound association, following directions and copying letters result in a readiness program designed to master these skills. They do not include items which reflect emergent literacy research (Day & Day, 1986). Instead, the tasks provide the teacher with fragmented data such as how well the child copies geometric forms

and matches initial sounds to representative pictures but not how well the child understands the reading process. Recent editions of these tests show that no significant alterations have been made to measure children's conscious awareness of the form, use and function of written language prior to formal literacy instruction.

The predictable value of standardized reading readiness tests has long been questioned by researchers. For example, Karlin (1971) summarizes various reports of the predictive validity of reading readiness tests and reports the correlations range from .40 - .60 with later reading achievement. Coltheart (1979) and Glazzard (1977) show that variables tapped by reading readiness tests are not predictively successful nor theoretically informative. Other researchers account for the variable predictive efficiency of such tests by noting that the variables, either predictor or criterion, have been conceptualized in very general or restrictive terms (Feshbach, Adelman & Fuller, 1977). Although it has been known for some time that the best predictors of reading achievement are those tests that most closely resemble tasks involved in reading (Karlin, 1971), schools continue to use group standardized measures to identify high-risk children, establish individual and group baseline information and make curriculum decisions (Hiebert, 1986).

In recent years efforts have increased toward the construction of more efficient prediction instruments. Specifically, instruments in which individual differences in acquiring emergent literacy

can be observed. A better understanding of emergent literacy has heightened awareness of the need for early identification of at-risk children as well as providing the means for fostering literacy.

Among the techniques which have contributed to our knowledge of emergent literacy behavior are the Concepts About Print Test (Clay, 1979), the Book Handling Task (Goodman & Altwerger, 1981), the Rhyme Reading Task (Morris, 1983), and the Metalinguistic Inventory (Evans, Taylor & Blum, 1979). While none shares a common task format, each of these instruments measures a discrete aspect of literacy behavior. All of these instruments employ a concrete stimulus for the child, examine print-related situations, and measure aspects of emergent literacy behaviors found to be related to reading success. The data from such instruments provide teachers with reliable information for grouping children, planning instruction, and reporting children's progress to parents and administrators. In spite of research results which show their effectiveness, informal assessment tasks are not commonly used as screening instruments at the preschool or primary levels.

The purpose of this investigation was to examine a number of literacy measures in light of their task demands, and their contribution to a composite picture of a child's literacy development. Answers to the following questions were sought:

1. Does an informal measure of print awareness, the Concepts About Print Test (CAP) estimate the level of reading achievement, (a) for first graders, (b) for high-risk primary grade students?

2. Does a standardized reading readiness battery, the Metropolitan Readiness Test (MRT) estimate the level of reading achievement, (a) for first graders, (b) for high-risk students?

3. Does the combination of effective predictors of literacy development differ (a) for first graders, (b) for the high-risk students?

Thus, the major focus was on comparing the effectiveness of informal and standardized readiness measures in assessing the literacy development of both normal first graders and high-risk, primary grade children.

Method

The Sample

For the purposes of this study, 87 subjects from primary grade classrooms were chosen--four classes of first grade students and three classes of high-risk students. Selection of both groups of subjects involved intact classrooms. The 60 first grade students were from a suburban school in upstate New York. The 27 high-risk primary grade children were from three intact classes of language-delayed students--one each of six-year-olds, seven-year-olds, and eight-year-olds from a suburban school serving only language-delayed children.

The Instruments

Three informal instruments were used to evaluate the literacy development of the subjects. These instruments included the Concepts About Print Test, the Writing Vocabulary Test, and the Sentence Dictation Test. In October, each of these instruments was administered individually to the first graders following the procedures outlined in Clay's The Early Detection of Reading Difficulties. The high-risk children were given the Concepts About Print Test and the Writing Vocabulary Test by one of the investigators. Scoring for each item was completed following the guidelines outlined by Clay.

The Concepts About Print Test, CAP, was selected for use in this study since it allowed the researchers to obtain information about the children's understanding of print concepts in a most efficient manner. A 24-item checklist of questions was asked while the storybook, Sand, was read to the child. Among the concepts assessed were those of letter, word, print direction, and uses of punctuation. The Writing Vocabulary Test was chosen as an inventory of the words each subject has under his/her control (spell correctly). This instrument consisted of an open-ended task in which the child was given ten minutes to write all the words s/he knew, starting with the child's name. As an evaluation of the child's ability to analyze and record the phonemes in individual words, the Sentence Dictation Test was administered. Two simple sentences were read to the subject, then repeated,

one word at a time, as the child attempted to write them. Each of these instruments has been normed and used with primary grade children. It was felt that since writing ability and reading ability both result from experiences with letters, words, and stories, the three tasks provided an opportunity for the child to show what s/he has learned about written language.

Readiness is commonly evaluated using a paper and pencil test. Thus, all subjects were given a formal assessment battery of readiness, the Metropolitan Readiness Tests. For the first grade subjects, the MRT, Level II, was group-administered by their classroom teachers in May of their kindergarten year. Subjects' scores were obtained from the school's printout. Following the guidelines for handicapped children, the MRT, Level I, was administered individually to each language-delayed subject by one of the researchers. Each high-risk subject's battery was hand-scored, following the directions in the MRT administration handbook. Local norms were established for this out-of-level test. In addition, each subject was also given the Peabody Picture Vocabulary Test as a measure of language capacity, a frequently used predictor of reading achievement.

For the first grade subjects, the Stanford Achievement Tests, Primary I, were administered by their classroom teachers in May. This battery was used as the measure of reading achievement. All test booklets were machine-scored and results for each subject were obtained from the school's printout. For the high-risk

subjects, there was no standardized reading battery administered in their school. The Letter Identification Test from Clay's Diagnostic Survey was administered to each language-delayed subject as a measure of reading ability. This task was administered by one of the researchers following the procedures outlined in Clay's Early Detection of Reading Difficulties.

Results

Relationship of Print Awareness to Reading Achievement

To determine the relationship between print awareness and reading achievement, the scores on the CAP were compared to the subtest scores and the total reading scores on the SAT using a Pearson-product moment correlation. For the first graders, the CAP was found to have significant correlations ($p<.001$) with the Word Recognition subtest (0.494), the Reading Comprehension subtest (0.512), the Word Study subtest (0.564), and the Total Reading score (0.531). For the high-risk subjects, their CAP scores were compared to the scores on the Letter Identification Test using a Pearson-product moment correlation. The CAP was found to have a significant ($p<.001$) with the Letter Identification Test (0.668). Thus, for both groups, normal and high-risk, the informal measure of print awareness, CAP, was found to estimate the level of reading achievement.

Relationship of Standardized Readiness Battery to Reading Achievement

Does a standardized reading readiness battery, the Metropolitan Readiness Tests, predict the level of reading achievement? Through

two different analyses, the answer to this question was found. Using the Pearson-product-moment correlation, the coefficients for the May administration of the MRT, Level II, (for first graders) with the SAT subtests were 0.570 for the Word Recognition subtest, 0.579 for Reading Comprehension, 0.564 for Word Study, and 0.554 for the Total Reading score. All correlations were significant at the .001 level. A linear regression analysis of MRT scores on the Total Reading scores was computed, resulting in an R-square equal to 0.306 ($F=5.148$, $p<.001$).

For the high-risk subjects, the results of the MRT, Level I, were compared to those of the Letter Identification Test using both a Pearson-product-moment correlation and a linear regression analysis. The Pearson-product-moment correlation for the MRT and the Letter Identification Test was 0.651 ($p<.001$). The linear regression analysis yielded an R-square equal to 0.423 ($F=4.285$, $p<.001$).

Comparison of Predictors in Estimating Achievement

To compare the effectiveness of each informal and formal instrument in estimating reading achievement, Pearson-product-moment correlations were calculated. For the first graders' SAT Total Reading scores, the strongest predictors were the Sentence Dictation Test, (0.710) and the Writing Vocabulary Test, (0.653). While the Sentence Dictation Test was also the strongest predictor for each of the three reading subtests, Word Recognition (0.709); Reading Comprehension (0.676); and Word Study (0.646), the Writing

Vocabulary Test was a strong predictor of the Word Study subtest, (0.584). The standardized readiness battery, the MRT, was the second strongest predictor for the Word Recognition subtest (0.570); and for the Reading Comprehension subtest (0.579). The Peabody Picture Vocabulary Test was not a significant predictor of any of the reading achievement subtests. The CAP showed significant correlations with the Total Reading and subtest scores, but was not as strong as the other informal instruments. (See Table 1.)

Insert Table 1 About Here

Next, multiple regression analyses were run to determine the effectiveness of different combinations of the informal literacy instruments in predicting the Total Reading scores for the first graders. The best combination of predictors was the Sentence Dictation Test and the Writing Vocabulary Test which resulted in an R-square of .528 ($F=39.21$, $p<.001$). The second best combination was the CAP and the Sentence Dictation Test with an R-square of .513 ($F=36.88$, $p<.001$). While the weakest combination was the CAP and the Writing Vocabulary with an R-square of .369 ($F=20.52$, $p<.001$), it was stronger than that of the six subtests that comprise the MRT battery (R-square = .306, $F=5.148$, $p<.001$). Thus, as predictors of first graders' reading achievement, the informal literacy measures were more effective than the formal readiness test battery.

Insert Table 2 About Here

Similarly, in order to establish the strongest predictors of reading achievement for the high-risk students, Pearson-product-moment correlations were calculated. (See Table 2.) For their Letter Recognition scores, the strongest predictors were the CAP (0.688), and the MRT (0.651). Next, multiple regression analyses were run to determine the effectiveness of different combinations of informal and formal instruments. The best combination of predictors was the CAP and the MRT which resulted in an R-square of 0.549 ($F=14.66$, $p<.001$). While the second best combination was the CAP and the Writing Vocabulary Test with an R-square of 0.481 ($F=11.16$, $p<.001$), it was a stronger predictor than the six subtests of the MRT battery (R-square equal to 0.423, $F=4.285$, $p<.001$). Therefore, for both groups of students the informal literacy measures yielded the best results. The Peabody Picture Vocabulary Test was not an effective predictor of reading achievement for either group. (See Tables 1 and 2.)

Discussion

While the effectiveness of the Concepts About Print Test for prediction of reading achievement in normal populations has been shown in past research (Day & Day, 1986; Freebody & Rust, 1985; Harlin, 1983), this is one of the first studies to support its efficiency as a predictor for high-risk children. In addition,

the combined predictive qualities of the CAP, the Sentence Dictation, the Writing Vocabulary Tests of Clay's Diagnostic Survey for first graders' reading achievement. While the standardized battery, the Metropolitan Readiness Tests, was found to be somewhat effective in identifying children at risk, its results were not sufficiently powerful to justify the time, effort, and expense of its administration. Therefore, this study's results support the use of informal, concrete tasks to assess the literacy development of both young and high-risk learners.

For teachers of primary grade children, as well as reading clinicians, the CAP has been shown to be an effective indicator of the child's knowledge and understanding of print concepts. Its ease of administration should recommend its use in the classroom as well as the reading clinic. The close correlation between the CAP and the measure of reading achievement, the SAT demonstrated the CAP's predictive qualities. As part of a preventative strategy, the CAP may be used to identify potential failure early in the school year, thus facilitating intervention strategies. For the reading clinician, the appropriateness of the CAP as a diagnostic tool for young disabled readers has been shown.

Although the program emphasis for high-risk children was different from that of normal first graders, in that it emphasized letter name knowledge, both the CAP and the Writing Vocabulary Test were sensitive to changes in their literacy development. These children are at a stage of literacy development similar

to the preschoolers studied by Mason (1982) who found that preschoolers acquired an increasing knowledge of letter names as they approached formal reading instruction. Thus, for this study, letter names were used as an indicator of print control.

The high correlations between the Sentence Dictation Test and the SAT show that not only is the informal task an accurate predictor of reading achievement, but also underscores the close relationship writing plays in reading acquisition. Like the CAP, this instrument is easy to administer and interpret. Within a ten to fifteen minute period, a teacher has acquired powerful data about the child's ability to encode written language, a skill that will impact on his reading success or failure.

The high correlations between the Writing Vocabulary Test and the Letter Identification Test support the contention by many researchers (Goodman & Goodman, 1983; Springate, 1983) that reading and writing are related tasks. Implications for including both "reading and writing" measures in pre-literacy assessments and program development for both normal and high-risk populations can be drawn from this information.

According to one theory of linguistic awareness, there is an interaction between reading acquisition and print awareness. As children learn to read, they become more sensitized to print (Ehri, 1979; Ryan, McNamera, & Kenney, 1977). While most of the children included in this study were not readers when pretested, the data revealed that they knew a great deal more than one would

expect about the functions of print and possessed a working knowledge of those functions as demonstrated by their performance on the writing tasks. This was true for the high-risk children who were not in a formal reading program, but who could write several words and name alphabet letters. Previous studies (Mason, 1980; Hiebert, 1979) revealed increasing reading readiness skills across normal preschool groups. The data from this study reveal a similar pattern for the high-risk group, but at a slower rate and more limited progression than in the normal population. While Mason (1982) found emergent literacy behaviors occurring naturally among normally developing preschoolers, VanKleeck and Schule (1987) suggest that emergent literacy behaviors do not develop naturally among language-delayed, high-risk children. Instead, they need active teaching both at home and school to foster the development of these concepts.

Observation of the data obtained from the informal measures, CAP and Writing Vocabulary, and the Metropolitan Readiness Test (MRT) reveal that while both are good predictors of letter name knowledge, the informal tests have several advantages.

The nature of the CAP measure allows the examiner to obtain individual profiles of children's print awareness. These profiles provide the teacher with usable information for instruction. For example, if a child does not know the left to right progression for reading, direct modeling and specific teaching can be incorporated in the child's program.

The Writing Vocabulary Test provides a measure of how children approach writing, their use of the alphabet and invented spelling patterns, and their formation of letters and letter sequences. Handicapped children should be offered the opportunity to develop a writing/reading relationship. Too often, these children are provided with oral instruction requiring oral feedback. Writing as a form of communication is often neglected because of predetermined notions that oral language and reading are precursors to writing. This relationship was not supported by the correlations of the Peabody Picture Vocabulary Test and the reading achievement measures.

For teachers, another advantage of informal measures is the opportunity to observe early emergence of oral/written language behaviors. In contrast, while administering group standardized reading readiness tests, which probe for the mastery of a skill, teachers have no indication of the strategies children are using to respond to those items. Informal measures provide a description of the emergent behavior and are more suitable indices for intervention needs. For example, the data from the CAP includes book handling tasks, basic concepts, e.g., front of book, first, last, etc., identification of print containing the message and reading terms such as letter and word. These data cannot be obtained from traditional standardized tests.

Rather than testing high-risk and normal children to determine if they are "ready" for formal reading, informal assessments should be periodically conducted to determine the extent to which

emergent literacy behaviors are developing. This diagnostic information can lead to meaningful intervention activities with print. No child should be deprived of print experiences. On the contrary, rich experiences with literature, shared reading, language experience stories, writing and reading simple messages should be the program emphasis.

References

Clay, M. (1979). The early detection of reading difficulty: A diagnostic survey with recovery procedures. Auckland, N.Z: Heinemann.

Coltheart, M. (1979). When can children learn to read - and when should they be taught? In T. G. Waller & G. E. MacKinnon (Eds.), Reading research - advances in theory and practice. New York: Academic Press.

Day, K. & Day, H. D. (1986). Tests of metalinguistic awareness. In D. Yaden & S. Templeton (Eds.), Metalinguistic awareness and beginning literacy: Conceptualizing what it means to read and write. Portsmouth, NH: Heinemann.

Ehri, L. (1979). Linguistic insights: Threshold of reading acquisition. In T. G. Waller and G. E. MacKinnon (Eds.), Reading Research: Advances in Theory Practice (Vol. 1). New York: Academic Press.

Evans, M., Taylor, N., & Blum, I. (1979). Children's written language awareness and its relation to reading acquisition. Journal of Reading Behavior, 11, 1-19.

Feshback, S., Adelman, H., & Fuller, W. (1977). Prediction of reading and related academic problems. Journal of Educational Psychology, 69, 299-308.

Freebody, P., & Rust, P. (1985). Predicting reading achievement in the first year of schooling: A comparison of readiness tests and instructional programs. Journal of School Psychology, 23, 144-155.

Glazzard, M. (1977). The effectiveness of three kindergarten predictors for first grade achievement. Journal of Learning Disabilities, 10, 95-99.

Goodman, K., & Goodman, Y. (1983). Reading and writing relationships: Pragmatic functions. Language Arts, 60, 590-599.

Goodman, Y., & Altwerger, B. (1981). Print awareness in preschool children: A working paper. Research paper #4, Occasional Papers, Program in Language and Literacy. Arizona Center for Research and Development, College of Education, University of Arizona.

Harlin, R. P. (1983). Primary children's print awareness as a predictor of reading achievement. Best Paper Award Monograph, New England Educational Research Organization.

Hiebert, E. (1979). The development of reading-related knowledge over the preschool years. Unpublished dessertation, University of Wisconsin, Madison.

Hiebert, E. H. (1986). Using environmental print in beginning reading instruction. In The Pursuit of Literacy: Early Reading and Writing, 73-80.

Karlin, R. (1971). Teaching elementary reading: Principles and strategies. New York: Harcourt, Brace, Jovanovich.

Mason, J. (1982). Acquisition of knowledge about reading: The preschool period (Technical Report No. 267). Urbana, IL: University of Illinois, Center for the Study of Reading.

Mason, J. (1980). When do children begin to read: An exploration of four-year-old children's letter and word reading competencies. Reading Research Quarterly, 15, 203-227.

Morris, D. (1983). Concept of word and phoneme awareness in the beginning reader. Research in the Teaching of English, 17, 359-373.

Ryan, E. B., McNamara, S. R., & Kenney, M. (1977). Linguistic awareness and reading performance among beginning readers. Journal of Reading Behavior, 9, 399-400.

Springate, K. (1983). Developmental trends and interrelationships among preprimary children's knowledge of writing and reading readiness skills. ED258178.

Van Kleeck, A., & Schuele, C. M. (1987). Precursors to literacy: Normal development. Topics in Language Disorders, 7, 13-44.

Table 1

Intercorrelations Between Predictor Variables for First Grade Subjects

	PPVT	CAP	DIC	WVC	SAT TOTAL	WORD REC	READ COMP	WORD STUDY
MRT	.381	.602	.748	.593	.561	.570	.579	.571
PPVT		.402	.275	.281	.193	.192	.226	.277
CAP			.650	.700	.531	.494	.512	.564
DIC				.741	.710	.709	.676	.646
WVC					.653	.523	.521	.584

Table 2

Intercorrelations Between Predictor Variables for High-Risk Subjects

	MRT	WVC	CAP	LET
PPVT	.202	.497	.471	.472
MRT		.443	.582	.651
WVC			.631	.568
CAP				.668